



**QUEEN'S  
UNIVERSITY  
BELFAST**

## **Synergy Between an Improved Covariate Unit Root Test and Cross-sectionally Dependent Panel Data Unit Root Tests, with two Applications**

Hadri, K., Kurozumi, E., & Yamazaki, D. (2015). Synergy Between an Improved Covariate Unit Root Test and Cross-sectionally Dependent Panel Data Unit Root Tests, with two Applications. *Manchester School*, 83(6), 676-700. <https://doi.org/10.1111/manc.12080>

**Published in:**  
Manchester School

**Document Version:**  
Peer reviewed version

**Queen's University Belfast - Research Portal:**  
[Link to publication record in Queen's University Belfast Research Portal](#)

### **Publisher rights**

© 2014 The University of Manchester and John Wiley & Sons

This is the peer reviewed version of the following article: Hadri, K., Kurozumi, E. and Yamazaki, D. (2014), Synergy between an Improved Covariate Unit Root Test and Cross-sectionally Dependent Panel Data Unit Root Tests. The Manchester School., which has been published in final form at doi: 10.1111/manc.12080

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving

### **General rights**

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### **Take down policy**

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact [openaccess@qub.ac.uk](mailto:openaccess@qub.ac.uk).

# Synergy Between an Improved Covariate Unit Root Test and Cross-sectionally Dependent Panel Data Unit Root Tests<sup>12</sup>

Kaddour Hadri

Eiji Kurozumi

Daisuke Yamazaki

Queen's University Management School  
Queen's University Belfast

Department of Economics  
Hitotsubashi University

Department of Economics  
Hitotsubashi University

November 22, 2013

Revised: April 30, 2014

## Abstract

This paper proposes the use of an improved covariate unit root test which exploits the cross-sectional dependence information when the panel data null hypothesis of a unit root is rejected. More explicitly, to increase the power of the test, we suggest the utilization of more than one covariate and offer several ways to select the "best" covariates from the set of potential covariates represented by the individuals in the panel. Monte Carlo simulations show that some of our methods work well compared to using only one covariate. Employing our methods, we investigate the Prebisch-Singer hypothesis for nine commodity prices. Our results show that this hypothesis holds for all but the price of petroleum.

*JEL classification:* C22, C23, C32

*Key words:* covariate unit root test, cross-sectional dependence in panel data, point optimal test, squared correlation, power

---

<sup>1</sup>We are grateful to two anonymous referees for useful comments. We are also grateful to participants at the 17th African Econometric Society meeting and seminar participants at National University of Singapore, Kyoto University, Queen's University Belfast, and Hitotsubashi University. All errors are our responsibility.

<sup>2</sup>This research was supported by JSPS and BA under the Japan-Britain Research Cooperative Program, by the Seimei Foundation, the Ministry of Education, Culture, Sports, Science and Technology under Grants-in-Aid No.23243038 and by the Global COE program of the Research Unit for Statistical and Empirical Analysis in Social Sciences, Hitotsubashi University. Address correspondence to Eiji Kurozumi, Department of Economics, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan; e-mail: kurozumi@stat.hit-u.ac.jp

## 1. Introduction

Testing for a unit root has a long history and its application in economics is well understood. A variety of univariate unit root tests have been proposed in the existing literature. However, these tests are generally criticized as having low power, particularly when the series are close to the unit root processes. To increase the power of univariate tests, panel data unit root tests have been proposed and developed (cf. Baltagi, 2008 and Breitung and Pesaran, 2008). A typical example is the use of such tests to investigate whether the purchasing power parity (PPP) hypothesis holds among OECD countries. Another aspect of using panel data is that, generally, the cross-sections are correlated. This is particularly true in macro panel data where  $T$  and  $N$  are large. O'Connell (1998), using a simulation, was the first to show that panel data unit root tests are distorted considerably when likely cross-sectional dependency is not considered. How to treat cross-sectional dependence has been investigated vigorously during the past decade. For example, Bai and Ng (2004) and Moon and Perron (2004) assumed a common factor structure to model strong cross-sectional dependence. They proposed a method of extracting common factors to make the panel data cross-sectionally uncorrelated, which then allows researchers to apply panel unit root tests such as the IPS test by Im, Pesaran, and Shin (2003), the Fisher test by Maddala and Wu (1999) and Choi (2001), and the inverse normal test by Choi (2001). In contrast, Pesaran (2007) proposed augmenting the regressions with the cross-sectional averages, while Chang (2002) and Chang and Song (2009) used nonlinear instrumental variables to mop up the cross-sectional dependence. With the development of these panel unit root (and also panel cointegration) tests, they have been used in empirical analysis. For example, the PPP hypothesis have been investigated by many papers such as Banerjee, Marcellino and Osbat (2005) and Westerlund and Edgerton (2008), while economic convergence has been focused on such as Evans and Karras (1996a, b), Romero-Ávila (2009) and Lin and Huang (2012). Westerlund (2008) tested for the Fisher effect using panel data, and Wagner (2008) investigated the environmental Kuznets Curve. See also Banerjee and Wagner (2009) for the theoretical overview and empirical examples.

While we may be able to overcome the problem of the low power of univariate unit root tests by making use of panel data, it has been pointed out in the literature (cf. Pesaran,

2012) that it is difficult to interpret the results when panel unit root tests reject the null hypothesis. This is because typical panel data tests reject the null of a unit root if some of the individuals are stationary and others have a unit root. Therefore, the rejection of the null hypothesis implies that not all individuals have a unit root, but we do not know which individuals are stationary. In this case, we may partition cross-sectional units into sub-groups and/or estimate the proportion of stationary units, as suggested by Pesaran (2012). We can eventually go back to the univariate tests, which we shall consider in this paper. See also Elliott and Pesavento (2006) for a discussion of the problem of the panel data approach.

At the same time, significant efforts have been devoted to improve the power of univariate tests. For instance, Hansen (1995) proposed augmenting the regression for the ADF test with covariates correlated with the disturbance term of the process on which we want to test the unit root hypothesis. This covariate ADF (CADF) test was further extended to a point-optimal covariate (POC) unit root test by Elliott and Jansson (2003). The power function of the POC test is tangent to the Gaussian power envelope at some point of the alternative. Juhl and Xiao (2003) proposed modifying the POC test by introducing the standard of optimality proposed by Cox and Hinkley (1974) to obtain the optimal point optimal covariate (OPOC) unit root test. More recently, Fossati (2013) extended the covariate unit root tests to models with structural breaks, while Westerlund (2013) allowed for conditional heteroskedasticity. These studies showed that the powers of the ADF and the ADF-GLS tests could be much improved if we can find covariates that are highly correlated with the disturbance term. In other words, the power improvement of these covariate unit root tests crucially depends on whether we can find appropriate covariates.

Although we may naturally find covariates in some cases, such as in the investigation of a technology shock by Christiano, Eichenbaum, and Vigfusson (2003), it is not always easy to find them in many practical situations. However, in cases in which the null hypothesis of a unit root is rejected when investigating macro panel data, one option is to test for a unit root in each cross-sectional unit using covariate tests. In this case, the natural candidates for covariates are the series of individuals other than the one being focused on, because macro panel data are typically cross-sectionally correlated. Although most empirical analyses have used only one covariate for covariate unit root tests, such as in Elliott and Pesavento

(2006), Amara and Papell (2006), and Christopoulos and León-Ledesma (2008), there is no justification to do so. In fact, we can expect that covariate tests with several covariates would be more powerful than those with only one covariate as long as we choose an appropriate set of covariates, as was considered by Lee and Tsong (2011).

In this study, we propose using the OPOC unit root test that exploits the cross-sectional dependence information contained in panel data when the null hypothesis of a unit root is rejected by panel unit root tests. We develop several selection rules to help us to choose appropriate covariates from those available. In addition to the factor model approach proposed by Lee and Tsong (2011), we propose two other procedures. One is based on the asymptotic power functions and the other one on the adjusted long-run squared correlation. We will show that the latter two methods work reasonably well in finite samples, while the factor model approach has a problem in terms of both size and power.

The rest of this paper is organized as follows. Section 2 considers the case in which the null hypothesis of panel unit roots is rejected and explains how to proceed with the OPOC unit root test in such a case. We propose three selection rules for covariates in Section 3 and investigate their finite sample properties in Section 4. Our methods are applied to the Prebisch-Singer hypothesis in Section 5. Concluding remarks are given in Section 6. Finally, the OPOC test is explained in detail in the Appendix.

## 2. Univariate Unit Root Test Revisited

Let us consider the following panel model:

$$z_{it} = \beta_{i,0} + \beta_{i,1}t + u_{it} \quad \text{for } i = 1, \dots, N \quad \text{and} \quad t = 1, \dots, T. \quad (1)$$

We call model (1) the trend case and the case with no linear trend ( $\beta_{i,1} = 0$  for all  $i$ ) the constant case. We would like to know if  $z_{it}$  are (trend) stationary or have a unit root.

Suppose that panel unit root tests are implemented on  $z_{it}$  and the unit root null hypothesis is rejected. In this case, it is difficult to interpret the result of the tests because panel unit root tests typically reject the null hypothesis even when only some of the cross-sectional units are stationary. That is, the rejection by the panel unit root tests implies that at least some of the individuals are stationary, but they do not indicate which individuals are stationary.

When the null hypothesis of a unit root is rejected by panel unit root tests, the tests are often then implemented on sub-groups of the cross-sectional units and/or the proportion of stationary units is estimated as suggested by Pesaran (2012). Eventually, we need to test for a unit root for each individual. That is, we need to rely on univariate unit root tests.<sup>3</sup>

Suppose, without loss of generality, that we now want to know if the first variable,  $y_t \equiv z_{1t}$ , is a unit root process. In this case, the common practice is to consider a univariate model for  $y_t$  given by

$$\Delta y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \psi_1 \Delta y_{t-1} + \cdots + \psi_p \Delta y_{t-p} + u_{y,t} \quad (2)$$

and to test for the null hypothesis of  $\rho = 0$  against the alternative of  $\rho < 0$ . Many unit root tests have been proposed, including the ADF test by Dickey and Fuller (1979) and Said and Dickey (1984) and the ADF-GLS test by Elliott, Rothenberg, and Stock (1996). However, in this study, we focus on a version of a covariate unit root test by Hansen (1995) because of its relative high power. Hansen (1995) proposed using the covariates  $x_t$ , which are  $I(0)$  variables<sup>4</sup> and are correlated with  $u_{y,t}$ , to improve the power of the ADF tests, and considered augmenting (2) with  $x_t$ , as follows:

$$\Delta y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \psi_1 \Delta y_{t-1} + \cdots + \psi_p \Delta y_{t-p} + \gamma' x_t + e_{y,t}. \quad (3)$$

The test is based on the  $t$ -statistic for  $\rho$  in (3). The improvement of power comes from the fact that the covariate vector  $x_t$  is correlated with  $u_{y,t}$ . Here, part of the fluctuation in  $u_{y,t}$  is explained by  $x_t$  and thus the variance of  $e_{y,t}$  in (3) becomes smaller than that of  $u_{y,t}$ . As a result, we can estimate  $\rho$  more efficiently with (3) than with (2). Hansen's covariate unit root test has been further refined by Elliott and Jansson (2003) who proposed a point optimal covariate (POC) unit root test by considering the local-to-unity system where  $\rho = -c/T$  for  $c \geq 0$ . One of the important characteristics of this test is that the critical values and the

---

<sup>3</sup>We could implement multivariate unit root tests, as described by Fountis and Dickey (1989), Shin (2004), and Ahlgren and Nyblom (2008) among others, or the cointegration tests described by Johansen (1991, 1995) on small sub-groups of cross-sectional units. However, the critical values of these tests depend on the true number of unit root processes in the groups. In addition, these tests tend to inherit the problem of low power from the univariate unit root tests. For these reasons, we do not pursue multivariate tests in this study.

<sup>4</sup>The power would be improved even if the covariates are  $I(1)$  variables. However, in this case, the critical values will depend on nuisance parameter as well as the number of  $I(1)$  variables. We then do not pursue the case of  $I(1)$  covariates.

power of the test depend on the so called long-run squared correlation  $R^2$ , which is defined in the Appendix.<sup>5</sup> The test becomes more powerful as  $R^2$  takes larger values. Finally, Juhl and Xiao (2003) extended it to the optimal point optimal covariate (OPOC) unit root test (see the Appendix for details on these tests). In the following discussion, we focus on the OPOC test.

To implement the OPOC test, we first need to find candidates for covariates. In this study, we consider the series of individuals other than the one we are focusing on as potential covariates  $(z_{2t}, \dots, z_{Nt})$  because macro panel data are typically (strongly) cross-sectionally correlated. However, we need to be careful when using these variables as covariates, because covariates must be  $I(0)$  variables. Note that we are considering the case in which we reject the null hypothesis of a unit root for panel data,  $z_{1t}, \dots, z_{Nt}$ , so some are (trend) stationary, but others may still have a unit root. The problem is that we do not know which variables are stationary, and yet covariates must be stationary, so the argument becomes circular. To avoid the circularity problem, we propose first to apply univariate unit root tests such as the ADF-GLS test and stationarity tests such the KPSS test by Kwiatkowski, Phillips, Schmidt and Shin (1992) to  $z_{2t}, \dots, z_{Nt}$ . If  $z_{it}$  is determined to be stationary, then we consider  $z_{it}$  in level as a covariate while the first differenced series  $\Delta z_{it} = z_{it} - z_{it-1}$  should be used when  $z_{it}$  has a unit root. It is sometimes the case that the  $I(0)/I(1)$  nature is inconclusive from standard univariate tests because of their low power. This typically happens when  $z_{it}$  is strongly serially correlated. In this case, we take the first difference of  $z_{it}$  as suggested by Hansen (1995).

In summary, when we reject the null hypothesis of a unit root for panel data, we test for a unit root for each individual using the OPOC test. In this case, if  $y_t = z_{1t}$  is of interest, then the candidates for covariates are  $z_{2t}, \dots, z_{Nt}$  but they should be first-differenced unless they are determined to be  $I(0)$  variables by standard univariate/stationarity tests.

### 3. Selection of Covariates

Once we find the candidates for covariates, we next need to choose appropriate covariates

---

<sup>5</sup>The long-run squared correlation  $R^2$  is unknown in practice and we have to estimate it. Accordingly, the critical values are determined based on the estimated  $R^2$  in empirical analysis.

that are used in the OPOC test. Although only one covariate has been used in most of previous studies, it would be better to use more covariates to increase the power of the test. In fact, it is known that the power of the OPOC test increases as  $R^2$  gets larger and that  $R^2$  becomes larger as the number of covariates increases. This implies that, at least theoretically, we should use as many covariates as possible. However, this is not necessarily a good strategy, for two reasons. The first is that, in finite samples, increasing the number of covariates means decreasing the degrees of freedom in a sample, which may result in a loss of power. In addition, using too many covariates results in an unstable estimation of the model because of the lower degrees of freedom, which may cause the test to collapse. In fact, our preliminary simulations show that the test with too many covariates always rejects the null hypothesis, even if  $y_t$  has a unit root. For these reasons, we need apposite selection rules for covariates.<sup>6</sup> Therefore, in this section, we propose three different approaches, all of which takes into account the degrees of freedom.

The first method we propose is to make use of the asymptotic local power of the OPOC test, which is defined as the asymptotic power of the test when the true value of parameter  $\rho$  is characterized as local to unity such that  $\rho = -c/T$  for  $c \geq 0$ . Suppose that the true value of  $c$  is equal to  $c^*$ . In this case, the asymptotic power apparently depends on  $c^*$ . We also know that the power of the OPOC test increases as  $R^2$  gets larger as explained in the previous section. In addition, as explained in the appendix, we need to pre-specify the specific value of  $c$  denoted as  $\bar{c}$ , to construct the OPOC test statistic. As a result, the asymptotic power function depends on  $\bar{c}$  as well as  $c^*$  and  $R^2$ , and so we denote it as  $h_{poc}(c^*, \bar{c}, R^2)$ . Note that in general, the asymptotic power can be seen as an approximation of the finite sample power, and is used to investigate the theoretical property of the test. It should be noted that the main difference between the POC and OPOC tests is in the value of  $\bar{c}$  used for the test statistic.

Now suppose we have a set of covariates with  $R^2 = R_1^2$ , which is assumed to be known

---

<sup>6</sup>It may be possible to make use of the existing information criteria such as AIC and BIC. However, most of the information criteria are designed to satisfy their own standards of optimality such as the minimization of the Kullback-Leibler information and the maximization of the posterior probability, which are not necessarily related with the problem we have in this paper. Because we are considering directly taking into account the trade-off between the power gain with covariates and the power loss by the low degrees of freedom, we do not pursue information criteria in this paper.



for the time being, and that the total number of regressors is  $k_1$ . In this case, the effective sample size is  $T - k_1$ , and  $\rho^*$  can be expressed as  $\rho^* = -c^*/T = -c_1/(T - k_1)$ , where  $c_1 = c^*(T - k_1)/T$ . Then, the corresponding asymptotic power against  $\rho = \rho^*$  becomes  $p_1 = h_{poc}(c_1, \bar{c}_1, R_1^2)$ , where  $\bar{c}_1$  is the pre-specified value of  $\bar{c}$  required to construct the OPOC test and depends on the value of  $R_1^2$ . We adjust the effective sample size because we expect it to better approximate the finite sample power than does the asymptotic power without the adjustment of the degrees of freedom. This will be investigated in the next section. Similarly, if we use another set of covariates with a known  $R^2 = R_2^2$  and the total number of regressors is equal to  $k_2$ , then the asymptotic power is given by  $p_2 = h_{poc}(c_2, \bar{c}_2, R_2^2)$ , where  $c_2 = c^*(T - k_2)/T$ . For these two sets of covariates, we choose the first set (the second set) if  $p_1 > p_2$  ( $p_1 < p_2$ ). The key feature of this procedure is that even if  $R_2^2 > R_1^2$ , it is possible for  $p_1$  to be greater than  $p_2$ , so we prefer the first set of covariates. This is illustrated in Figure 1. In the latter, even though the power function for  $R_2^2$  dominates that for  $R_1^2$ , if we use too many covariates to attain  $R^2 = R_2^2$ , then the effective sample size decreases. As a result,  $p_2$ , the corresponding asymptotic power against  $\rho = \rho^*$ , may then be smaller than  $p_1$ .

In practice, we do not know the true values of  $R_1^2$  and  $R_2^2$  and then we have to use the estimates of them,  $\hat{R}_1^2$  and  $\hat{R}_2^2$ , and  $\bar{c}_1$  and  $\bar{c}_2$  are determined based on  $\hat{R}_1^2$  and  $\hat{R}_2^2$ . As a result, the asymptotic powers also depend on these estimates and are denoted as  $\hat{p}_1$  and  $\hat{p}_2$ , respectively. Even in this case, it is possible for  $\hat{p}_1$  to be greater than  $\hat{p}_2$  even if  $\hat{R}_1^2 < \hat{R}_2^2$ .

In this procedure, we have to choose a specific value of  $c^*$  as a benchmark. Note that if  $c^*$  is too small or too large, then the difference between the power functions for different values of  $R^2$  is small, and we may obtain similar results for any set of covariates. Thus, we propose choosing a value of  $c^*$  that maximizes the difference between the power functions for  $R^2 = 0$  and 0.9. According to our calculation, we obtain that  $c^* = 4.0$  in the constant case and  $c^* = 5.8$  in the trend case. We call this selection procedure the adjusted power rule.

The second selection rule mimics the well-known adjusted squared correlation in the usual sense. That is, we define

$$\hat{\hat{R}}^2 = 1 - \frac{T-1}{T-k}(1 - \hat{R}^2),$$

where  $k$  is the total number of regressors and  $\hat{R}^2$  is the estimated long-run square correlation.

We propose choosing the set of covariates that attains the highest  $\hat{R}^2$ . Although this is an ad hoc rule with no theoretical support, it is the easiest rule of the three to apply in an empirical analysis. Note that the term  $(T - 1)/(T - k)$  can be interpreted as a penalty on using  $k$  covariates. The penalty becomes heavier as we use more covariates. The validity of this penalty is investigated in the next section. We refer to this selection procedure as the adjusted long-run square correlation rule.

The third method of choosing covariates is basically the same as that of Lee and Tsong (2011). Here, a factor model is assumed for a set of variables. Because the common factors play a key role in cross-sectional dependence, these factors are the natural possible candidates for covariates. We thus estimate the common factors using the principal component method proposed by Bai (2003) and use the estimated common factors as covariates. Note that Bai's estimation requires that all the variables should be stationary. Because we are testing for a unit root in  $y_t$ , we should take the first difference of  $y_t$  to guarantee stationarity, while  $z_{2t}, \dots, z_{Nt}$  are used in the level or first differenced, according to the result of univariate pre-tests conducted for covariates. The advantage of this method is that it is computationally easy to obtain the covariates even if  $N$  is relatively large. However, there is no guarantee that the factor structure is the correct specification or that the long-run squared correlation will be highest when using the selected covariates. We call this selection procedure the common factor rule.

## 4. Finite Sample Properties of the Selection Rules

### 4.1. The Effect of the Finite Sample Adjustment

We first investigate how well the finite sample adjustments proposed in the previous section work in order to assess the effect of the decrease in the degrees of freedom. In particular, we investigate the performance of the adjusted power rule and the adjusted  $R^2$  rule when the number of covariates,  $k$ , increases while the true value of  $R^2$  is fixed for any value of  $k$ . In this case, the finite sample power decreases as  $k$  gets larger because the degrees of freedom decreases while  $R^2$  is fixed. That is, we examine how well our two selection rules mimic the decrease in power caused by using too many covariates.

The data we consider is generated by the following:

$$\begin{aligned} z_{1t} &= \beta_{1,0} + \beta_{1,1}t + u_{1t}, \quad u_{1t} = \phi u_{1t-1} + \varepsilon_{1t}, \\ z_{it} &= \beta_{i,0} + \beta_{i,1}t + \varepsilon_{it}, \quad (i = 2, \dots, N) \end{aligned} \quad (4)$$

for  $t = 1, \dots, T$  where  $\beta_{j,0} = \beta_{j,1} = 0$  for  $j = 1, \dots, N$  and  $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt}]' \sim i.i.d.N(0, \Sigma)$  with

$$\Sigma = \begin{bmatrix} 1 & R & 0 & \cdots & 0 \\ R & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad (5)$$

In this case,  $z_{2t}$  is the only effective covariate and the other variables,  $z_{3t}, \dots, z_{Nt}$ , do not help to increase the power. The autoregressive parameter,  $\phi = 1 - c/T$ , is set with  $c = 2, 4$ , and  $6$  for the constant case, and  $c = 4, 6$ , and  $8$  for the trend case. We construct the OPOC test statistic for  $z_{1t}$  with the number of covariates ranging from 1 to 15 where  $z_{2t}$  is always included as a covariate, and the other variables,  $z_{3t}, z_{4t}, \dots$ , are added as the number of covariates increases. This implies that the long-run squared correlation is always equal to  $R^2$ , from the structure of the covariance matrix,  $\Sigma$ . We choose  $R^2 = 0.5$  while  $T = 100$  and  $200$ . The number of replications is 5,000 and all computations are conducted using the GAUSS matrix language.<sup>7</sup>

Figures 2(i-a) and (i-b) (the first column) show the finite sample power of the OPOC test (lines with “+”) and the theoretical asymptotic local power obtained using the effective sample size (lines with “o”) when  $T = 200$ .<sup>8</sup> These are drawn as functions of the number of covariates. Since the slopes of the powers are important in assessing the usefulness of the proposed selection rule, the level of the asymptotic power function is adjusted so that the average of the powers at 15 points in the figure becomes the same as that of the finite sample powers. From Figure 2, we can see that, as expected, the finite sample power of the test decreases as the number of the (invalid) covariates increases. In addition, the theoretical power based on our adjusted power rule mimics the decrease in the finite sample power well,

<sup>7</sup>The GAUSS code is available upon request.

<sup>8</sup>We obtained a similar result for  $T = 100$  and thus omit it to save space.

although the former is slightly steeper than the latter. This implies that the adjusted power rule imposes a slightly heavier penalty on the number of covariates than expected. Therefore, we can say that this rule is slightly conservative in that it tends to avoid using too many covariates.

Similarly, Figures 2(ii-a) and (ii-b) (the second column) shows the finite sample power (lines with “+”) and the adjusted long-run squared correlation  $\bar{R}^2$  (lines with “o”) when  $T = 200$ . Again,  $\bar{R}^2$  mimics the slope of the finite sample power but it is steeper than the nominal power. Therefore, the adjusted  $R^2$  rule also results in a conservative choice of covariates.

#### 4.2. The Performance of the Proposed Rules

We next investigate the finite sample properties of the three selection procedures proposed in this paper using Monte Carlo simulations. Since most of previous empirical work have only used one covariate, we compare the performance of the OPOC test with our selection rules with the test using only one covariate. In the latter test, the covariate is chosen to maximize the estimated  $R^2$ .

The data for the simulations is generated in the same way as (4), with two types of covariance matrix,  $\Sigma$ . In the first case (DGP1), we define  $\Sigma$  as

$$\Sigma = \begin{bmatrix} 1 & \sqrt{R^2/(N-1)} & \sqrt{R^2/(N-1)} & \cdots & \sqrt{R^2/(N-1)} \\ \sqrt{R^2/(N-1)} & 1 & 0 & \cdots & 0 \\ \sqrt{R^2/(N-1)} & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \sqrt{R^2/(N-1)} & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

In this case,  $R^2$  always increases as the number of covariates increases, and the theoretical long-run squared correlation reaches  $R^2$  when all the  $N - 1$  variables are used as covariates. With this setting, we can investigate how effective our selection rule is in increasing the power of the test.

In the second simulation case (DGP2), we consider (4) with the covariance matrix defined by (5). In this case, using only one covariate is the most efficient method, so we can assess the loss caused by using more than one covariate based on our rules.

In the simulations we set the cross-sectional dimension to  $N = 5, 10$ , and  $20$  and the time series dimension to  $T = 100$  and  $200$ . We set  $R^2 = 0.3, 0.6$ , and  $0.9$  as weakly, intermediately, and strongly dependent cases, respectively. The AR parameter  $\phi$  is set to  $1$  under the null hypothesis while  $\phi = 0.98, 0.96, 0.94, 0.92$  and  $0.90$  under the alternative. To focus on the performance of the selection rules, we assume that  $z_{2t}, \dots, z_{Nt}$  are known to be  $I(0)$ , so we use these variables in levels throughout the simulations. We calculate the OPOC tests using the three selection rules described previously, as well as the test using only one covariate, for the purpose of comparison. For the latter test, the covariate is chosen to maximize the estimated  $R^2$ . The number of replications is  $5,000$  with a significance level equal to  $0.05$ .

For the adjusted power, the  $\bar{R}^2$  rules, and the test with one covariate, we estimate  $R^2$  under the null hypothesis as suggested by Elliott and Jansson (2003), with the order of lag selected using the Bayesian information criterion and the maximum lag length set to  $4$ . When  $N = 5$  and  $10$ , we calculate the asymptotic power adjusted by the effective sample size based on  $\hat{R}^2$  (the estimated  $R^2$ ) and  $\hat{\bar{R}}^2$  for all the possible  $1$  to  $N - 1$  combinations of  $z_{2t}, \dots, z_{Nt}$ . We then choose a set of covariates for which these criteria are maximized. On the other hand, when  $N = 20$ , we restrict the maximum number of covariates (denoted as  $\bar{N}$ ) to  $9$  for  $T = 100$  and  $13$  for  $T = 200$  to avoid the test statistic collapsing because of a lack of degrees of freedom. In addition, even if the maximum number of covariates is restricted, the number of possible combinations is still too large to conduct simulations when  $N = 20$ . In this case, the covariates are not selected from all possible combinations, but are chosen sequentially, as follows: First, we choose one covariate for which the selection rules are maximized. We then select an additional covariate from the remainder so that the selection rules are maximized with two covariates. We proceed by adding covariates until the number of covariates reaches  $\bar{N}$ . Note that the number of covariates chosen in this procedure ranges from  $1$  to  $\bar{N}$ . Finally, we determine a set of covariates from the  $\bar{N}$  candidates so that the selection rules are maximized.

For the adjusted power rule, we need the theoretical power functions for different values of  $R^2$ . We calculate the asymptotic power functions for  $R^2 = 0, 0.1, 0.2, \dots, 0.9, 0.92, 0.94, 0.96, 0.98, 0.99$  and  $c = 0, 1, \dots, 15$ , and obtain the power corresponding to the given  $c_1$  and  $R^2$  by interpolation.

To apply the common factor rule, we have to determine the number of factors. As per de Silva, Hadri and Tremayne (2009), we adopt the Hannan-Quinn-type criterion  $HQ_4$ . This is a modified version of the information criteria proposed by Bai and Ng (2002), with the maximum number of factors set to 4 when  $N = 5$  and 10, and set to 8 when  $N = 20$ . Using this criterion, we extract common factors from  $\{\Delta z_{1t}, z_{2t}, \dots, z_{Nt}\}$  using the principal component method of Bai (2003).

Table 1 shows the empirical sizes of the tests in the constant case. We can see that the tests with one covariate ( $\Lambda_1$ ), with the adjusted power rule ( $\Lambda_{pow}$ ), and with the adjusted long-run squared correlation rule ( $\Lambda_{\bar{R}^2}$ ), are slightly over-sized when  $T = 100$  but the empirical sizes get closer to the nominal one when  $T = 200$ , whereas the test with the common factor rule ( $\Lambda_{com}$ ) suffers from severe size distortion when  $N = 5$ . In this case, the test rejects the null hypothesis in almost all cases.<sup>9</sup> However, the size of this test is close to the nominal one for larger values of  $N$ .

The numbers in parentheses in the last three columns are the average number of selected common factors and the average number of selected covariates. For the common factor rule, the selected number of common factors is almost always 4 when  $N = 5$  and 1 when  $N = 10$  and 20. On the other hand, the performances of the adjusted power rule and the  $\bar{R}^2$  rule are very similar and more covariates tend to be chosen for DGP1 than DGP2, as expected considering the structure of these DGPs.

Table 2 shows the corresponding values for the trend case, and the relative performance of the tests remains the same.

Figures 3(i-a)–(i-c) (the first column) show the size-adjusted power of the four tests for DGP1 with  $T = 200$  and  $R^2 = 0.6$  in the constant case (similar results are obtained for  $R^2 = 0.3$  and 0.9 and when  $T = 100$ , so we omit the results here to save space). Because the true  $R^2$  value increases as the number of covariates increases, the tests using our selection rules are expected to be more powerful than the test using only one covariate. In fact, from the figure, we observe that the advantage of the adjusted power rule and the  $\bar{R}^2$  rule over

---

<sup>9</sup>As pointed out by one of the referees, one of the possible reasons for this poor performance may be as follows: The number of common factors is determined to be 4 when  $N = 5$  and in this case, using these four estimated common factors as covariates in regression would departure seriously from the correlation structure of the data generating process.

the test using only one covariate is pronounced. In particular, when  $R^2$  is 0.9 (which is not reported in the figure), the difference in power can be more than 0.6 when  $N = 10$  and  $R^2 = 0.9$ . On the other hand, the test using the common factor rule is not as powerful as the other two rules, and performs poorly when  $N = 5$ .

In contrast to DGP1, the test using one covariate is most favorable in terms of power for DGP2, as shown in Figures 3(ii-a)–(ii-c) (the second column). This is because only one covariate is valid while the other variables are meaningless in this case. However, the power difference is relatively minor, particularly when  $N = 5$  and 10. When  $N = 20$  and  $R^2 = 0.3$ , the power difference becomes slightly larger but it is at most 0.1. For DGP2, the test using the common factor rule is the least favorable for all cases in terms of power.

Figure 4 correspond to the trend case for DGP1 and 2. Again, we observe similar performances to those obtained in the constant case.

In addition to the above two cases, we also consider as the data generating process the factor model given by

$$z_t = \beta_0 + \beta_1 t + u_t, \quad A(L)u_t(\rho) = \lambda f_t + \varepsilon_t,$$

where  $z_t = [z_{1t}, \dots, z_{Nt}]'$ ,  $u_t = [u_{1t}, \dots, u_{Nt}]'$ ,  $f_t \sim i.i.d.N(0, \sigma_f^2)$  is a one-dimensional common factor,  $\lambda = [\lambda_1, \lambda_c']'$ , where  $\lambda_c' = [\lambda_2, \dots, \lambda_N]'$  is an  $(N - 1)$ -dimensional loading vector and  $\varepsilon_t = [\varepsilon_{1t}, \dots, \varepsilon_{Nt}]' \sim i.i.d.N(0, I_n)$  is independent of  $f_t$ . In this model, the long-run squared correlation can be expressed as

$$R^2 = \frac{\sigma_f^4 \lambda_1^2 (\lambda_c' \lambda_c)}{(1 + \sigma_f^2 \lambda_1^2) \left\{ 1 + \sigma_f^2 (\lambda_c' \lambda_c) \right\}}.$$

In order to control the value of  $R^2$ , we set  $\lambda_1 = 1$  and  $\lambda_c' \lambda_c = N - 1$  and choose the value of  $\sigma_f^2$  so that  $R^2 = 0.3, 0.6$ , and  $0.9$ . More precisely, we first generate  $\lambda_i^* \sim U(0.5, 1.5)$  for  $i = 2, \dots, N$  independently and normalize them as  $\lambda_c = \sqrt{(N - 1)/(\lambda_c^{*'} \lambda_c^*)} \lambda_c^*$ , where  $\lambda_c^* = [\lambda_2^*, \dots, \lambda_N^*]'$ , so that the restriction  $\lambda_c' \lambda_c = N - 1$  holds. In this case,  $\sigma_f^2$  is the positive solution of the quadratic function of  $k$  given by  $(N - 1)(1 - R^2)k^2 - NR^2k - R^2 = 0$ .

In this case, the test with the common factor rule performs poorly when  $N = 5$  under the null hypothesis as in the case of DGP1 and 2, whereas it performs best in terms of power

when  $N = 10$  and  $20$ , although the difference in power among the tests with three rules proposed in this paper is not necessarily large. It is natural to obtain this result because the data generating process in this case includes the common factor (we omit the table and the figure to save space).

In summary, our simulations reveal that the two selection rules proposed in this study, the adjusted power rule and the adjusted long-run squared correlation rule, perform relatively well when compared to the test using only one covariate, which has often been used in previous studies. In particular, the advantage of these two tests in terms of power is much more pronounced when the true GDP is favorable to our rules, whereas the loss when using them in the least favourable case (only one covariate is relevant) is relatively small.

## 5. Empirical Applications: Prebisch-Singer hypothesis

The Prebisch-Singer (PS) hypothesis states that relative commodity prices follow a downward secular trend. Prebisch (1950) and Singer (1950) claimed that there had been a downward long-term trend in these relative prices and that the decline in these prices was likely to continue. The main theoretical explanations given for this negative long-term trend are: (a) income elasticities of demand for primary commodities are lower than those for manufactured commodities; (b) an absence of differentiation among commodity producers leading to highly competitive markets; (c) productivity differentials between North and South; (d) asymmetric market structures: the presence of oligopolistic rents for the North and zero economic profit for competitive commodity producers in the South; (e) the inability of wages to grow in the presence of an “unlimited” supply of labor at the subsistence wage in primary commodity-producing countries (Lewis, 1954); and (f) a decline in demand from industrial countries. However, this effect has recently decreased as a result of the growing demand from emerging market countries, such as China, India, and Brazil. The consequences of this hypothesis are very important for developing countries because many of them depend on only a few primary commodities to generate most of their export earnings. If we assume that  $y_{it}$ , the relative commodity price  $i$ , is generated by a stationary process around a time trend ( $I(0)$ ), then

$$y_{it} = \beta_{i,0} + \beta_{i,1}t + u_{it}, \quad t = 1, \dots, T, \quad (6)$$



where the random variable  $u_{it}$  is stationary with mean 0 and variance  $\sigma_{i,u}^2$ . The parameter of interest is the slope  $\beta_{i,1}$ , which is predicted to be negative under the PS hypothesis. However, if commodity prices are generated by a so called difference-stationary (DS or I(1)) model, which would imply that  $y_{it}$  is non-stationary, then

$$\Delta y_{it} = \beta_{i1} + \zeta_{it}, \quad t = 1, \dots, T, \quad (7)$$

where  $\zeta_{it}$  is stationary. It is now well known that if  $y_{it}$  is an I(1) process, then using equation (6) to test the null hypothesis  $\beta_{i,1} = 0$  will result in severe size distortions. This, in turn, lead to the null being wrongly rejected when no trend is present, even asymptotically. Alternatively, if the true generating process is given by equation (6) and we base our test on equation (7), then our test becomes inefficient and less powerful than the one based on the correct equation. Therefore, when testing the PS hypothesis we must first test the order of integration of our relative commodity prices to ensure we use the correct equation.

In this subsection, we investigate the I(0)/I(1) properties of nine real commodity prices (zinc, tin, oil, wool, iron, aluminum, beef, coffee, and cocoa) relative to the US CPI index using annual data from 1960 to 2007. We first treat the data set as panel data and apply the PANIC test proposed by Bai and Ng (2004). We assume a common factor structure in  $u_{i,t}$ , such that  $u_{it} = \lambda_i f_t + \varepsilon_{it}$ , where  $f_t$  is an  $r$ -dimensional common factor,  $\lambda_{it}$  is a  $1 \times r$  loading vector, and  $\varepsilon_{it}$  is an idiosyncratic error. We estimate  $\varepsilon_{it}$  using a principal component analysis, then apply the Fisher test and the inverse normal test to the estimated  $\varepsilon_{it}$ . The results are shown in Panel (a) in Table 3. The number of common factors is estimated to be 4 by the  $HQ_4$  proposed by de Silva, Hadri, and Tremayne (2009). Both tests reject the null hypothesis of a unit root for the idiosyncratic errors using the size-adjusted critical values. We also apply the panel trend stationarity tests  $ZA_{spc}$  and  $ZA_{la}$  proposed by Hadri and Kurozumi (2012) and the test by Harris, Leybourne, and McCabe (2005). The results of the tests are consistent and imply that some of the prices can be characterized as trend stationary processes. However, they do not tell us which prices are trend stationary.

We next conduct univariate tests. We test for the null hypothesis of a unit root for each price using the ADF-GLS test, with the lag length selected by the modified AIC of Ng and Perron (2001), while the null of trend stationarity is checked using the bias-corrected

version of the KPSS test, with the boundary condition equal to 0.95. This version of the test was developed by Kurozumi and Tanaka (2010) by correcting the bias in the test statistic previously proposed by Kwiatkowski, Phillips, Schmidt, and Shin (1992). The results are given in the second and third columns in Panel (b) in Table 3. We find strong evidence of stationarity for the prices of wool and aluminum. The bias-corrected version of the KPSS test rejects the null of trend stationarity for the price of tin at the 5% significance level and for the prices of petroleum, tin, and beef at the 10% significance level. However, we need to interpret these results carefully, because the KPSS test is known to suffer from size distortions when a process is strongly serially correlated.

We next apply the OPOC test with the adjusted power rule to each of the prices other than wool and aluminum, which we have already established to be  $I(0)$ . Because the covariates must be stationary, we take the first difference of these prices when using these variables as covariates. If we reject the null of a unit root for some of the prices, then we treat those variables as trend stationary, use them in levels as covariates, and test the other variables again. We repeat the procedure until we cannot find additional evidence of stationarity. The results are given in Panel (b) in Table 3. The fourth column reports the number of covariates chosen by our selection rule, while the fifth column reports the estimated long-run squared correlation when those covariates are used for testing. From the results given in the sixth column, we can reject the null of a unit root for five prices: zinc, tin, petroleum, iron, and coffee. In addition to the prices of wool and aluminum for which we have already rejected the null of a unit root, we find that seven of the nine commodities have trend-stationary prices. However, they might be very persistent, as shown, *inter alia*, by Cuddington and Jerret (2008). For the prices of beef and cocoa, we cannot reach a conclusive result. Note that the ADF-GLS test rejects the unit root hypothesis for only two of nine prices while the OPOC test can reject this hypothesis for additional five series, which implies the usefulness of the OPOC test in view of power. Based on these results, we test the PB hypothesis based on equation (6) for the seven trend-stationary prices, and estimate the slope for beef and cocoa using (7). The final column reports the  $p$ -values of the one-sided tests based on the  $t$ -statistics calculated using the autocorrelation-heteroskedasticity consistent standard errors using the quadratic spectral kernel. Here, the bandwidth was selected using the method

proposed by Andrews (1991). We can see, with the exception of the price of petroleum, which has a positive coefficient, the estimates of the slope coefficients are all significant and negative. Therefore, we find strong evidence of the PB hypothesis for seven commodity prices and weak evidence for cocoa. The exception is the price of petroleum.

We also investigated the same data using a shorter sample period from 1960 to 2002, because a structural change might have occurred in the early 2000s, as pointed out by Arezki, Hadri, Kurozumi, and Rao (2012). However, the results of the tests were similar to the above case and we reached the same conclusion.

## 6. Concluding Remarks

In this paper, we investigate an improved covariate unit root tests. We propose to use the OPOC test and exploit the cross-sectional dependence information to ascertain which cross-section variables are stationary after that the panel data null hypothesis of a unit root is rejected. We also suggest the use of more than one covariate and propose in this respect three selection rules for choosing potential covariates. The Monte Carlo simulations show that the rules based on adjusted power and the adjusted  $R^2$  work reasonably well, but that the common factor rule must be used with caution.

In empirical work the choice of covariates should not be confined only by the ones obtained from panel data with cross-sectional dependency. If we can find other covariates, then it would be better to include these series when applying our selection rules. Moreover, in such a situation, we may be able to consider a panel version of the covariate unit root test as described by Chang and Song (2009) and Westerlund (2012). Either way, our results show that combining covariate tests and panel tests could be helpful in empirical applications.

## Appendix: Optimal Point Optimal Covariate Unit Root Test

Suppose we want to know if the first variable  $z_{1t}$  has a unit root. To focus on  $z_{1t}$ , let  $y_t = z_{1t}$  and  $x_t = [z_{2t}, \dots, z_{Nt}]'$ . Stacking variables in the cross-sectional direction, we then consider the following model:

$$z_t = \beta_0 + \beta_1 t + u_t, \quad A(L)u_t(\phi) = \varepsilon_t, \quad \text{where} \quad u_t(\phi) = \begin{bmatrix} (1 - \phi L)u_{y,t} \\ u_{x,t} \end{bmatrix}, \quad (8)$$

$z_t = [y_t, x_t']'$ ,  $\beta_0 = [\beta_{y,0}, \beta_{x,0}']'$ ,  $\beta_1 = [\beta_y, \beta_{x,1}']'$ ,  $u_t = [u_{y,t}, u_{x,t}']'$ ,  $\varepsilon_t = [\varepsilon_{y,t}, \varepsilon_{x,t}']'$ , and  $A(L)$  is a lag polynomial of order  $p$  with  $L$  being the lag operator. Since  $A(L)$  is supposed to be invertible by Assumption A1 below,  $u_t(\phi)$  is assumed to be stationary. Note that the variables and parameters are decomposed conformably with  $z_t = [y_t, x_t']'$ . We define the long-run variance of  $u_t(\phi)$  and the long-run squared correlation as

$$\Omega = A^{-1}(1)\Sigma A'^{-1}(1) = \begin{bmatrix} \omega_{yy} & \omega_{yx} \\ \omega_{xy} & \Omega_{xx} \end{bmatrix} \quad \text{and} \quad R^2 = \omega_{yy}^{-1}\omega_{yx}\Omega_{xx}^{-1}\omega_{xy}, \quad \text{respectively.}$$

Model (8) allows for heterogeneity and cross-sectional dependence in  $u_t$  through the lag-polynomial  $A(L)$  and the innovation variance matrix  $\Sigma$ . Note that the factor structure is also included as a special case by assuming that  $A(L)u_t(\phi) = \varepsilon_t$  with  $\varepsilon_t = \Lambda f_t + e_t$ , where  $f_t$  is an  $r$ -dimensional common factor,  $\Lambda$  is an  $N \times r$  loading matrix, and  $e_t$  consists of the idiosyncratic errors.

**Assumption A1** (a)  $\{\varepsilon_t\}$  is a martingale difference sequence with respect to  $\mathcal{F}_t = \sigma\{\varepsilon_t, \varepsilon_{t-1}, \dots\}$  with  $E[\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}] = \Sigma > 0$  for all  $t$ . (b)  $\sup_t E\|\varepsilon_t\|^{2+\kappa} < \infty$  for some  $\kappa > 0$ . (c)  $|A(z)| = 0$  implies  $|z| > 1$ . (d)  $u_0, u_{-1}, \dots, u_{-p}$  are  $O_p(1)$  and independent of  $T$ .

Since we are interested in whether  $y_t$  is a unit root process, we consider the following testing problem:

$$H_0 : \phi = 1 \quad \text{vs.} \quad H_1 : |\phi| < 1.$$

Elliott and Jansson (2003) proposed a point optimal covariate unit root (POC) test by considering the local-to-unity system for (8). More precisely, they assumed that  $\phi = 1 - c/T$  for  $c \geq 0$  and proposed constructing the likelihood ratio (LR) test statistic,  $\Lambda(1, \bar{\phi})$ , by assuming that  $c = \bar{c}$  (or  $\bar{\phi} = 1 - \bar{c}/T$ ). This test has been shown to depend on only  $c$ ,  $\bar{c}$ , and  $R^2$  asymptotically, and the asymptotic local power function can then be written as  $h_{poc}(c, \bar{c}, R^2)$ . By the Neyman-Pearson lemma, the LR test is a most powerful test against  $c = \bar{c}$  under the assumption of normality. The Gaussian power envelope, which was also investigated by Hansen (1995), is then given by  $h_{poc}(c, c, R^2)$ . We can see that the power function of the LR test is tangent to the power envelope at  $c = \bar{c}$ , but is generally lower than the envelope at  $c \neq \bar{c}$ . Note that we need to pre-specify  $\bar{c}$  to construct the test statistic  $\Lambda(1, \bar{\phi})$ . Elliott and Jansson (2003) recommended using  $\bar{c} = 7$  in the constant case and  $\bar{c} = 13.5$  in the trend case.

Although Elliott and Jansson (2003) showed that the power function of the POC test is close to the power envelope in a wide range of alternatives for different values of  $R^2$ , Juhl and Xiao (2003) pointed out that there are other possibilities for the choice of  $\bar{c}$ . This is because the value of  $\bar{c}$  suggested by Elliott and Jansson (2003) is based on the choice of Elliott, Rothenberg, and Stock (1996), which implies that the power function of the POC test using the suggested  $\bar{c}$  is tangent to the 50% point of the power envelope only when  $R^2 = 0$ , so that  $h_{poc}(\bar{c}, \bar{c}, 0) = 0.5$ . However, because the power function depends on the true value of  $R^2$ , Juhl and Xiao (2003) concluded that the choice of  $\bar{c}$  should also depend on  $R^2$ . Moreover, there is no theoretical reason to choose the value of  $\bar{c}$  at which the power function is tangent to the 50% point of the power envelope. Rather, they proposed choosing  $\bar{c}$  for which

$$\int_0^\infty [h_{poc}(c, c, R^2) - h_{poc}(c, \bar{c}, R^2)] dc \quad (9)$$

is minimized; that is, for a given value of  $R^2$ , the average loss of power compared to the power envelope is minimized at  $\bar{c}$ . This criterion of optimality was originally proposed by Cox and Hinkley (1974) and also adopted by Kurozumi (2003) in a different situation. Juhl and Xiao (2003) called this POC test, in which  $\bar{c}$  minimizes (9), the optimal point optimal covariate (OPOC) unit root test. The optimal values of  $\bar{c}$  for a given value of  $R^2$  are given in Table 1 in Juhl and Xiao (2003). Therefore, the main difference between POC and OPOC is in the calculation of  $\bar{c}$ .

Following Elliott and Jansson (2003), the OPOC test is constructed as follows. First, we estimate  $\beta = [\beta'_0, \beta'_1]'$  from the quasi-differenced series under the null and alternative hypotheses, that is,

$$\tilde{\beta}(r) = \left[ S \left( \sum_{t=1}^T d_t(r) \hat{\Omega}^{-1} d'_t(r) \right) S \right]^{-1} \left[ S \sum_{t=1}^T d_t(r) \hat{\Omega}^{-1} z_t(r) \right]$$

for  $r = 1$  and  $r = \bar{\phi}$ , where  $z_1(r) = [y_1, x'_1]'$  for  $t = 1$  and  $z_t(r) = [(1 - rL)y_t, x'_t]'$  for  $t > 1$ ,  $d'_1(r) = [I_N, I_N]$  for  $t = 1$ , and

$$d'_t(r) = \begin{bmatrix} 1 - r & 0 & (1 - rL)t & 0 \\ 0 & I_{N-1} & 0 & tI_{N-1} \end{bmatrix}$$

for  $t > 1$ ,  $S = \text{diag}\{I_N, 0\}$  in the constant case and  $S = I_{2N}$  in the trend case,  $B^-$  is the Moore-Penrose inverse of a matrix  $B$ , and  $\hat{\Omega}$  is the estimator of the long-run variance  $\Omega$  under

the null hypothesis. We next construct the detrended series given by

$$\tilde{u}_t(r) = z_t(r) - d'_t(r)\tilde{\beta}(r)$$

for  $r = 1$  and  $\bar{\phi}$ . Using  $\tilde{u}_t(r)$ , we estimate the VAR model of order  $p$  and obtain the estimated residual  $\tilde{\varepsilon}_t(r)$  and the estimator of variance  $\tilde{\Sigma}(r)$ . Then, the test statistic is given by

$$\Lambda(1, \bar{\phi}) = T \left[ \text{tr} \left( \tilde{\Sigma}^{-1}(1) \tilde{\Sigma}(\bar{\phi}) \right) - (N - 1 + \bar{\phi}) \right].$$

The rejection region is the left-hand tail of the distribution of  $\Lambda(1, \bar{\phi})$ .

## References

- [1] Ahlgren, N. and J. Nyblom (2008). Tests against Stationary and Explosive Alternatives in Vector Autoregressive Models. *Journal of Time Series Analysis* 29, 421-443.
- [2] Amara, J. and D. H. Papell (2006). Testing for Purchasing Power Parity Using Stationary Covariates. *Applied Financial Economics* 16, 29-39.
- [3] Andrews, D. W. K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica* 59, 817-858.
- [4] Arezki, R., K. Hadri, E. Kurozumi and Y. Rao (2012). Breaking the Prebisch Singer Hypothesis Using Panel Data Stationarity Tests. Forthcoming in *Economics Letters*.
- [5] Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica* 71, 135-171.
- [6] Bai, J. and S. Ng (2002). Determining the Number of Factors in Approximate Factor Models. *Econometrica* 70, 191-221.
- [7] Bai, J. and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration. *Econometrica* 72, 1127-1177.

- [8] Baltagi, B. H. (2008). *Econometric Analysis of Panel Data* . Chichester, Wiley.
- [9] Banerjee, A., M. Marcellino and C. Osbat (2005). Testing for PPP: Should We Use Panel Methods? *Empirical Economics* 30, 77-91.
- [10] Banerjee, A. and M. Wagner (2009). Panel Methods to Test for Unit Roots and Cointegration, in T. C. Mills and K. Patterson ed., *Palgrave Handbook of Econometrics, Vol.2*. Palgrave Macmillan, New York.
- [11] Breitung, J. and Pesaran, M. H. (2008). Unit roots and cointegration in panels. In: L. Matyas and P. Sevestre (eds): *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*. Dordrecht: Kluwer Academic Publishers, 3rd edition, Chap. 9, pp. 279-322.
- [12] Chang, Y. (2002). Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. *Journal of Econometrics* 110, 261-292.
- [13] Chang, Y. and W. Song (2009). Testing for Unit Roots in Small Panels with Short-Run and Long-Run Cross-Sectional Dependencies *Review of Economic Studies* 76, 903-935.
- [14] Choi, I. (2001). Unit Root Tests for Panel Data. *Journal of International Money and Banking* 20, 249-272.
- [15] Christiano, L. J., M. Eichenbaum and R. Vigfusson (2003). What Happens After a Technology Shock? *NBER working paper* 769, 1-52.
- [16] Christopoulos, D. K. and M. A. León-Ledesma (2008). Time-Series Output Convergence Tests and Stationary Covariates. *Economics Letters* 101, 297-299.
- [17] Cox, D. R. and D. V. Hinkley (1974). *Theoretical Statistics*. Chapman and Hall, London.
- [18] Cuddington, J. and D. Jerret (2008). Super Cycles in Metal Prices? *IMF staff Papers* 55(4), 541-565.
- [19] de Silva, S., Hadri, K. and Tremayne, A. R. (2009). Panel Unit Root Tests in the Presence of Cross-Sectional Dependence: Finite Sample Performance and an Application. *Econometrics Journal* 12, 340-366.

- [20] Dickey, D. A. and W. A. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association* 74, 427-431.
- [21] Elliott, G. and M. Jansson (2003). Testing for Unit Roots with Stationary Covariates. *Journal of Econometrics* 115, 75-89.
- [22] Elliott, G. and E. Pesavento (2006). On the Failure of Purchasing Power Parity for Bilateral Exchange Rates after 1973. *Journal of Money, Credit and Banking* 38, 1405-1430.
- [23] Elliott, G., T. J. Rothenberg and J. Stock (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica* 64, 813-836.
- [24] Evans, P. and G. Karras (1996a). Convergence Revisited. *Journal of Monetary Economics* 37, 249-265.
- [25] Evans, P. and G. Karras (1996b). Do Economies Converge? Evidence from a Panel of U.S. States. *Review of Economics and Statistics* 78, 384-388.
- [26] Fountis, N. G. and D. A. Dickey (1989). Unit Root Nonstationarity in Multivariate Autoregressive Time Series. *Annals of Statistics* 17, 419-428.
- [27] Fossati, S. (2013). Unit Root Testing with Stationary Covariates and a Structural Break in the Trend Function. *Journal of Time Series Analysis* 34, 368-384.
- [28] Hadri, K. and E. Kurozumi (2012). A Simple Panel Stationarity Test in the Presence of Serial Correlation and a Common Factor. *Economics Letters* 115, 31-34.
- [29] Hansen, B. E. (1995). Rethinking the Univariate Approach to Unit Root Testing. *Econometric Theory* 11, 1148-1171.
- [30] Harris, D., S. Leybourne and B. McCabe (2005). Panel Stationarity Tests for Purchasing Power Parity with Cross-Sectional Dependence. *Journal of Business and Economic Statistics* 23, 395-409.



- [31] Im, K., M. H. Pesaran and Y. Shin (2003). Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics* 115, 53-74.
- [32] Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59, 1551-1580.
- [33] Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.
- [34] Juhl, T. and Z. Xiao (2003). Power Functions and Envelopes for Unit Root Tests. *Econometric Theory* 19, 240-253.
- [35] Kurozumi, E. (2003). Some Properties of the Point Optimal Invariant Test for the Constancy of Parameters. *Journal of the Japan Statistical Society* 33, 169-180.
- [36] Kurozumi, E. and S. Tanaka (2010). Reducing the Size Distortion of the KPSS Test. *Journal of Time Series Analysis* 31, 415-426.
- [37] Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics* 54, 159-178.
- [38] Lee, C. F. and C. C. Tsong (2011). Covariate Selection for Testing Purchasing Power Parity. *Applied Economics* 43:15, 1923-1933.
- [39] Lewis, A. (1954). Economic Development with Unlimited Supplies of Labour. *Manchester School of Economics and Social Studies* 22, 139-191.
- [40] Lin, P.-C. and H.-C. Huang (2012). Convergence in Income Inequality? Evidence from Panel Unit Root Tests with Structural Breaks. *Empirical Economics* 43, 153-174.
- [41] Maddala, G. S. and S. Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics* 61, 631-52.
- [42] Moon, R. and B. Perron (2004). Testing for Unit Roots in Panels with Dynamic Factors. *Journal of Econometrics* 122, 81-126.

- [43] Ng, S. and P. Perron (2001). Lag Length Selection and the Construction of Unit Root Tests With Good Size and Power. *Econometrica* 69, 1519-1554.
- [44] O'Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics* 44, 1-19.
- [45] Pesaran, M. H. (2007). A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence. *Journal of Applied Econometrics* 22, 265-312.
- [46] Pesaran, M. H. (2012). On the Interpretation of Panel Unit Root Tests. *Economics Letters*, forthcoming.
- [47] Prebisch, R. (1950). The Economic Development of Latin America and Its Principal Problems. *Economic Bulletin for Latin America* 7, 1-12.
- [48] Romero-Ávila, D. (2009). The Convergence Hypothesis for OECD Countries Reconsidered: Panel Data Evidence with Multiple Breaks, 1870-2003. *Manchester School* 77, 552-574.
- [49] Said, E. S. and D. A. Dickey (1984). Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order. *Biometrika* 71, 599-607.
- [50] Shin, K. (2004). A Multivariate Unit Root Test Based on the Modified Weighted Symmetric Estimator for VAR( $p$ ). *Journal of Applied Statistics* 31, 587-596.
- [51] Singer, H. (1950). The Distribution of Gains between Investing and Borrowing Countries. *American Economic review, Papers and Proceedings* 40, 473-485.
- [52] Wagner, M. (2008). The Carbon Kuznets Curve: A Cloudy Picture Emitted by Bad Econometrics? *Resource and Energy Economics* 30, 388-408.
- [53] Westerlund, J. (2008). Panel Cointegration Tests of the Fisher Effect. *Journal of Applied Econometrics* 23, 193-233.
- [54] Westerlund, J. (2012). Rethinking the Univariate Approach to Panel Unit Root Testing: Using Covariates to Resolve the Incidental Trends Problem. manuscript, Deakin University.

- [55] Westerlund, J. (2013). A Computationally Convenient Unit Root Test with Covariates, Conditional Heteroskedasticity and Efficient Detrending. *Journal of Time Series Analysis* 34, 477-495.
- [56] Westerlund, J. and D. L. Edgerton (2008). A Simple Test for Cointegration in Dependent Panels with Structural Breaks. *Oxford Bulletin of Economics and Statistics* 70, 665-704.

Table 1: Empirical size (constant case)

(a) DGP1

$T$	$N$	$R^2$	$\Lambda_1$	$\Lambda_{com}$	$\Lambda_{pow}$	$\Lambda_{\bar{R}^2}$
100	5	0.3	0.087	0.999 (4.0)	0.082 (3.2)	0.083 (3.4)
		0.6	0.077	1.000 (4.0)	0.072 (3.9)	0.073 (3.9)
		0.9	0.069	1.000 (4.0)	0.063 (4.0)	0.063 (4.0)
	10	0.3	0.089	0.069 (1.0)	0.070 (6.3)	0.072 (6.7)
		0.6	0.080	0.059 (1.0)	0.062 (8.0)	0.065 (8.1)
		0.9	0.077	0.058 (1.0)	0.066 (9.0)	0.066 (9.0)
	20	0.3	0.087	0.072 (1.0)	0.080 (7.2)	0.082 (7.4)
		0.6	0.083	0.061 (1.0)	0.070 (7.8)	0.070 (7.8)
		0.9	0.086	0.057 (1.0)	0.074 (8.2)	0.074 (8.2)
200	5	0.3	0.068	1.000 (4.0)	0.067 (3.8)	0.067 (3.8)
		0.6	0.062	1.000 (4.0)	0.058 (4.0)	0.058 (4.0)
		0.9	0.061	1.000 (4.0)	0.055 (4.0)	0.055 (4.0)
	10	0.3	0.070	0.051 (1.0)	0.065 (7.0)	0.065 (7.5)
		0.6	0.070	0.050 (1.0)	0.060 (8.8)	0.061 (8.8)
		0.9	0.068	0.053 (1.0)	0.060 (9.0)	0.060 (9.0)
	20	0.3	0.073	0.058 (1.0)	0.074 (10.7)	0.075 (11.1)
		0.6	0.070	0.052 (1.0)	0.068 (12.0)	0.068 (11.9)
		0.9	0.070	0.053 (1.0)	0.066 (12.6)	0.065 (12.5)

(b) DGP2

$T$	$N$	$R^2$	$\Lambda_1$	$\Lambda_{com}$	$\Lambda_{pow}$	$\Lambda_{\bar{R}^2}$
100	5	0.3	0.081	0.999 (4.0)	0.080 (2.0)	0.080 (2.4)
		0.6	0.065	1.000 (4.0)	0.066 (2.1)	0.066 (2.5)
		0.9	0.056	1.000 (4.0)	0.057 (2.0)	0.059 (2.5)
	10	0.3	0.074	0.069 (1.0)	0.072 (4.7)	0.073 (5.4)
		0.6	0.056	0.059 (1.0)	0.060 (4.7)	0.064 (5.6)
		0.9	0.051	0.057 (1.0)	0.061 (3.8)	0.065 (5.2)
	20	0.3	0.075	0.071 (1.0)	0.079 (6.8)	0.083 (7.5)
		0.6	0.060	0.058 (1.0)	0.076 (7.2)	0.080 (8.1)
		0.9	0.055	0.051 (1.0)	0.079 (7.0)	0.083 (8.7)
200	5	0.3	0.062	1.000 (4.0)	0.062 (2.0)	0.063 (2.4)
		0.6	0.054	1.000 (4.0)	0.055 (2.0)	0.055 (2.5)
		0.9	0.051	1.000 (4.0)	0.050 (2.0)	0.051 (2.5)
	10	0.3	0.066	0.052 (1.0)	0.066 (3.9)	0.064 (4.8)
		0.6	0.057	0.050 (1.0)	0.056 (3.7)	0.056 (5.0)
		0.9	0.056	0.050 (1.0)	0.055 (3.7)	0.055 (5.1)
	20	0.3	0.059	0.060 (1.0)	0.064 (8.1)	0.066 (9.8)
		0.6	0.050	0.053 (1.0)	0.062 (8.3)	0.061 (10.3)
		0.9	0.050	0.052 (1.0)	0.058 (7.3)	0.064 (10.3)

Note: The entries on the columns of  $\Lambda_1$ ,  $\Lambda_{com}$ ,  $\Lambda_{pow}$  and  $\Lambda_{\bar{R}^2}$  are the empirical sizes of the tests with only one covariate, the common factor rule, the adjusted power rule, and the adjusted long-run square correlation rule, respectively. The entries in parentheses on the column  $\Lambda_{com}$  is the average number of selected common factors in 5,000 replications, while those in parentheses on the columns  $\Lambda_{pow}$  and  $\Lambda_{\bar{R}^2}$  are the average number of covariates selected by these rules.

Table 2: Empirical size (trend case)

(a) DGP1

$T$	$N$	$R^2$	$\Lambda_1$	$\Lambda_{com}$	$\Lambda_{pow}$	$\Lambda_{\bar{R}^2}$
100	5	0.3	0.064	0.999 (4.0)	0.067 (3.4)	0.067 (3.4)
		0.6	0.058	1.000 (4.0)	0.065 (3.9)	0.066 (3.9)
		0.9	0.055	1.000 (4.0)	0.061 (4.0)	0.061 (4.0)
	10	0.3	0.064	0.051 (1.0)	0.066 (6.3)	0.067 (6.4)
		0.6	0.061	0.050 (1.0)	0.062 (8.1)	0.063 (8.0)
		0.9	0.056	0.053 (1.0)	0.068 (9.0)	0.068 (9.0)
	20	0.3	0.066	0.058 (1.0)	0.074 (6.4)	0.075 (6.4)
		0.6	0.066	0.054 (1.0)	0.066 (7.1)	0.066 (7.0)
		0.9	0.064	0.053 (1.0)	0.069 (7.6)	0.070 (7.5)
200	5	0.3	0.048	1.000 (4.0)	0.051 (3.8)	0.051 (3.8)
		0.6	0.047	1.000 (4.0)	0.053 (4.0)	0.053 (4.0)
		0.9	0.042	1.000 (4.0)	0.056 (4.0)	0.056 (4.0)
	10	0.3	0.060	0.048 (1.0)	0.058 (7.5)	0.058 (7.6)
		0.6	0.060	0.048 (1.0)	0.059 (8.8)	0.059 (8.8)
		0.9	0.056	0.047 (1.0)	0.057 (9.0)	0.057 (9.0)
	20	0.3	0.059	0.054 (1.0)	0.066 (10.1)	0.066 (10.1)
		0.6	0.058	0.050 (1.0)	0.062 (11.2)	0.062 (11.1)
		0.9	0.057	0.050 (1.0)	0.058 (12.1)	0.058 (12.0)

(b) DGP2

$T$	$N$	$R^2$	$\Lambda_1$	$\Lambda_{com}$	$\Lambda_{pow}$	$\Lambda_{\bar{R}^2}$
100	5	0.3	0.058	0.999 (4.0)	0.066 (2.3)	0.066 (2.4)
		0.6	0.057	1.000 (4.0)	0.062 (2.2)	0.063 (2.5)
		0.9	0.057	1.000 (4.0)	0.058 (2.1)	0.059 (2.5)
	10	0.3	0.058	0.055 (1.0)	0.058 (5.1)	0.059 (5.3)
		0.6	0.052	0.049 (1.0)	0.061 (5.2)	0.061 (5.7)
		0.9	0.054	0.051 (1.0)	0.061 (4.3)	0.062 (5.5)
	20	0.3	0.060	0.056 (1.0)	0.075 (6.5)	0.078 (6.6)
		0.6	0.058	0.057 (1.0)	0.076 (7.1)	0.076 (7.5)
		0.9	0.058	0.052 (1.0)	0.076 (7.1)	0.083 (8.6)
200	5	0.3	0.051	1.000 (4.0)	0.051 (2.3)	0.051 (2.4)
		0.6	0.052	1.000 (4.0)	0.053 (2.2)	0.053 (2.5)
		0.9	0.051	1.000 (4.0)	0.050 (2.1)	0.051 (2.5)
	10	0.3	0.049	0.047 (1.0)	0.058 (4.8)	0.057 (5.0)
		0.6	0.053	0.047 (1.0)	0.060 (4.2)	0.061 (5.0)
		0.9	0.054	0.049 (1.0)	0.058 (3.9)	0.060 (5.1)
	20	0.3	0.051	0.054 (1.0)	0.066 (8.8)	0.068 (9.4)
		0.6	0.052	0.053 (1.0)	0.068 (9.0)	0.069 (10.1)
		0.9	0.054	0.048 (1.0)	0.063 (7.9)	0.067 (10.4)

Note: The entries on the columns of  $\Lambda_1$ ,  $\Lambda_{com}$ ,  $\Lambda_{pow}$  and  $\Lambda_{\bar{R}^2}$  are the empirical sizes of the tests with only one covariate, the common factor rule, the adjusted power rule, and the adjusted long-run square correlation rule, respectively. The entries in parentheses on the column  $\Lambda_{com}$  is the average number of selected common factors in 5,000 replications, while those in parentheses on the columns  $\Lambda_{pow}$  and  $\Lambda_{\bar{R}^2}$  are the average number of covariates selected by these rules.

Table 3: Prebisch-Singer hypothesis

## (a) Panel unit root and stationarity tests

estimated # of the common factors:	4
PANIC test for the idiosyncratic errors	
(Fisher test)	33.783**
(inverse normal test)	-2.821***
panel stationarity tests	
( $ZA_{spc}$ )	-1.892
( $ZA_{la}$ )	-2.518
(HLM test)	-0.707

## (b) univariate unit root and stationarity tests

	ADF-GLS	KPSS	OPOC test			$t_{\beta_{i1}}$ p-value
			# of cov.	$R^2$	test stat.	
zinc	-2.416	0.070	6	0.779	-0.143**	0.000
tin	-1.573	0.148**	2	0.954	-6.905**	0.000
petro.	-1.675	0.129*	4	0.981	-0.300**	1.000
wool	-3.467***	0.076	-	-	-	0.000
iron	-1.384	0.136*	5	0.862	-2.306**	0.000
aluminum	-3.905***	0.061	-	-	-	0.000
beef	-1.536	0.122*	5	0.998	7.797	0.000
coffee	-2.495	0.113	2	0.919	-3.388**	0.000
cocoa	-2.082	0.112	4	0.979	-0.273	0.058

Note: Rejections at 10%, 5%, and 1% significance level are denoted by \*, \*\*, and \*\*\*, respectively. The ADF-GLS and the OPOC tests suppose the null hypothesis of a unit root against the alternative of trend stationarity, so that the rejection by these tests implies that the series is trend stationary. On the other hand, the KPSS test is designed for the null hypothesis of stationarity against the alternative of a unit root and then the rejection supports the unit root hypothesis.

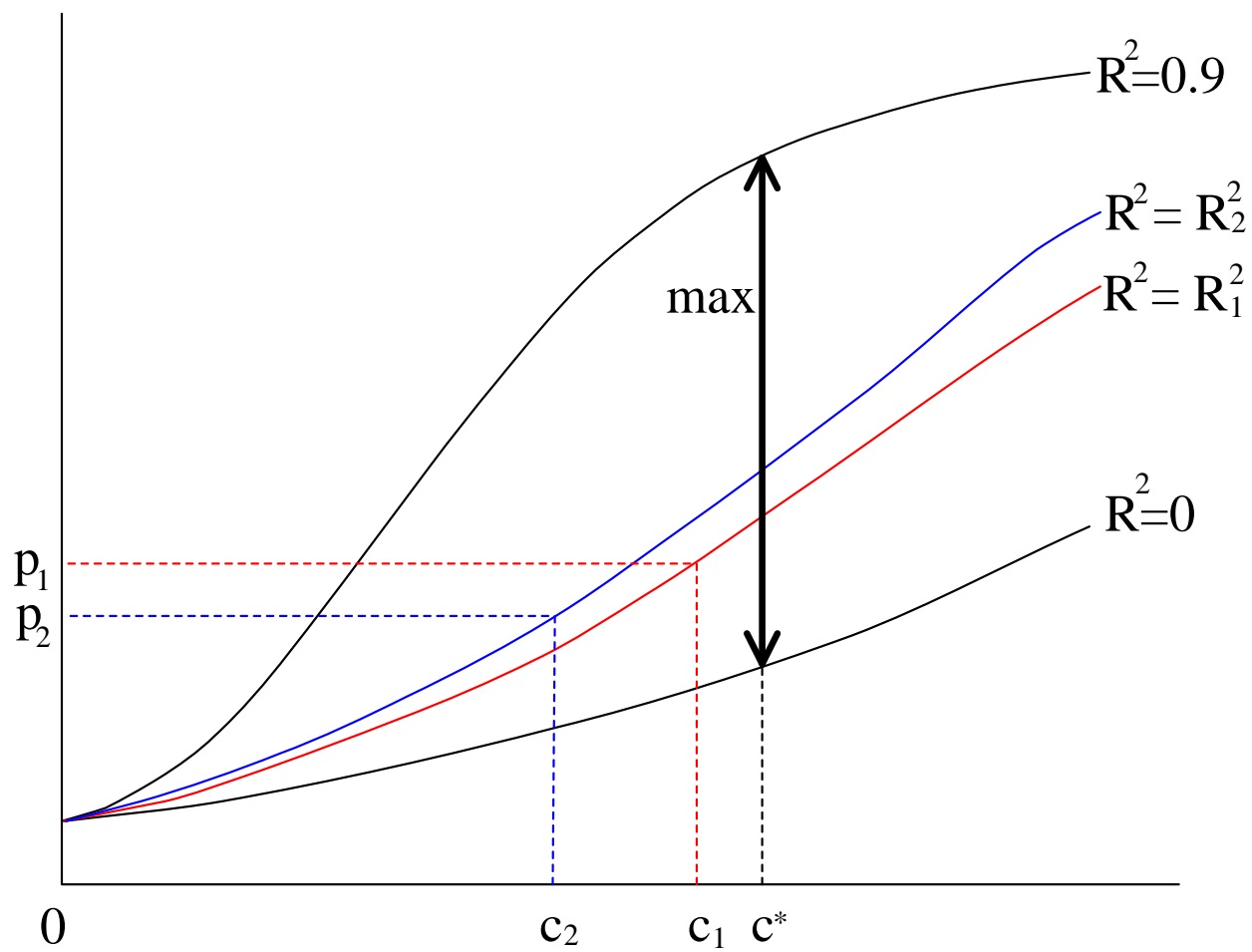
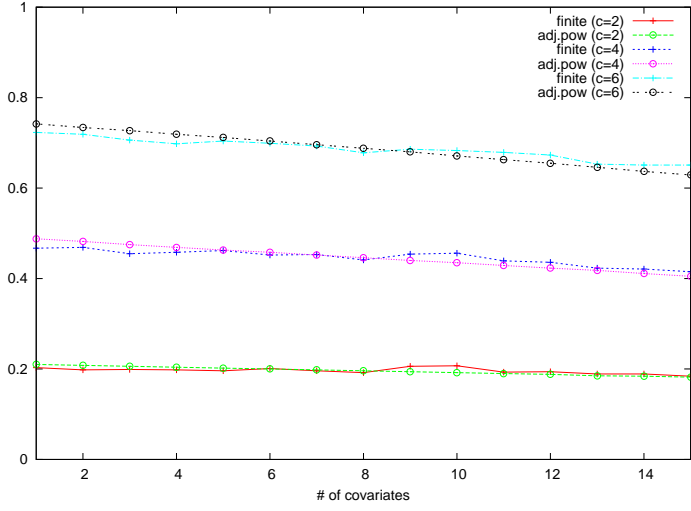
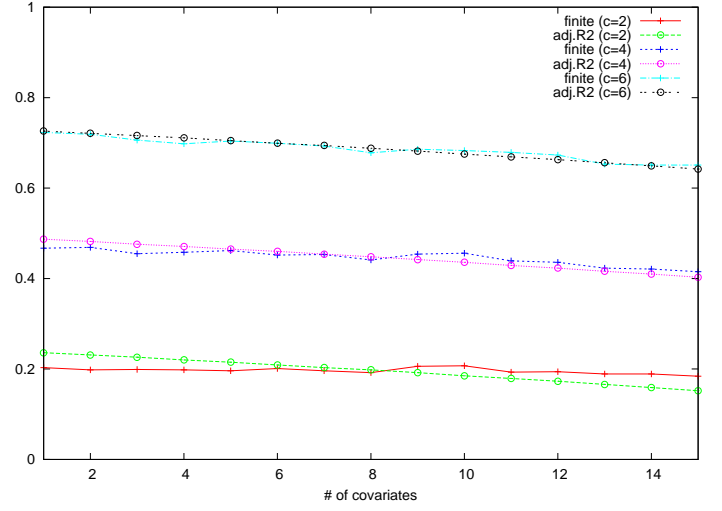


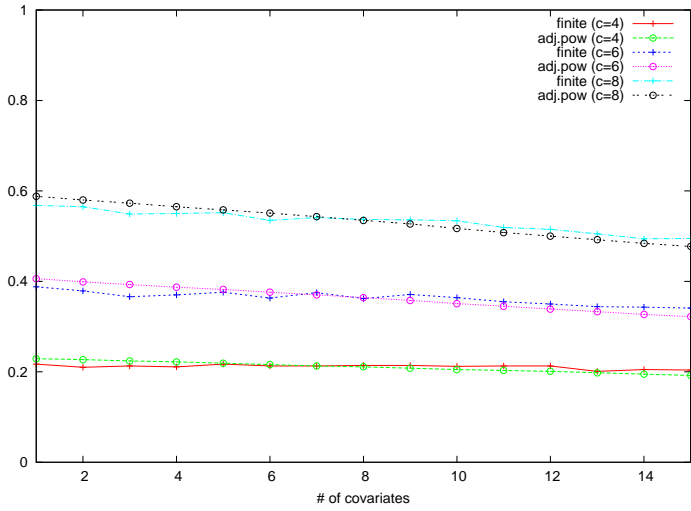
Figure 1: The selection rule of covariates



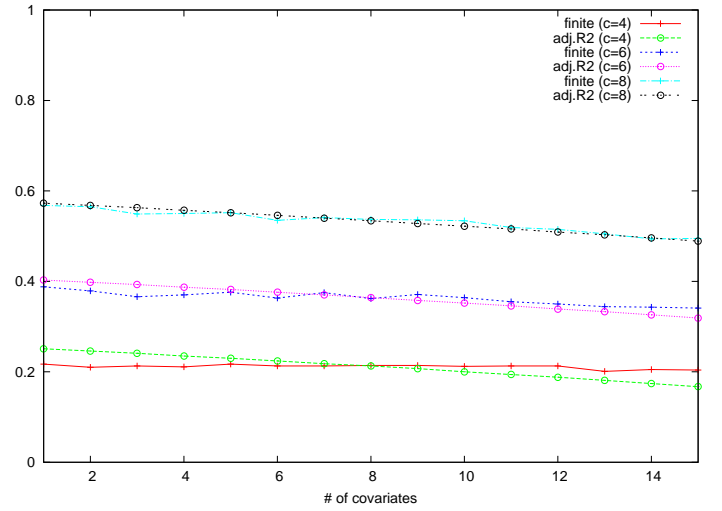
(i-a) constant case (adjusted power)



(ii-a) constant case ( $\bar{R}^2$ )



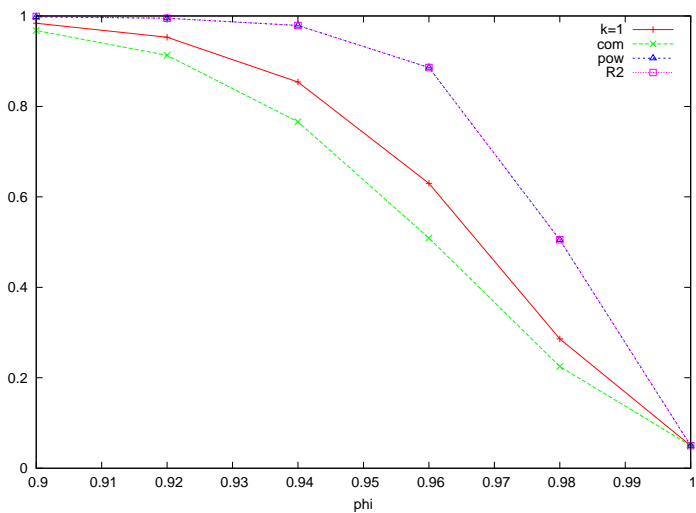
(i-b) trend case (adjusted power)



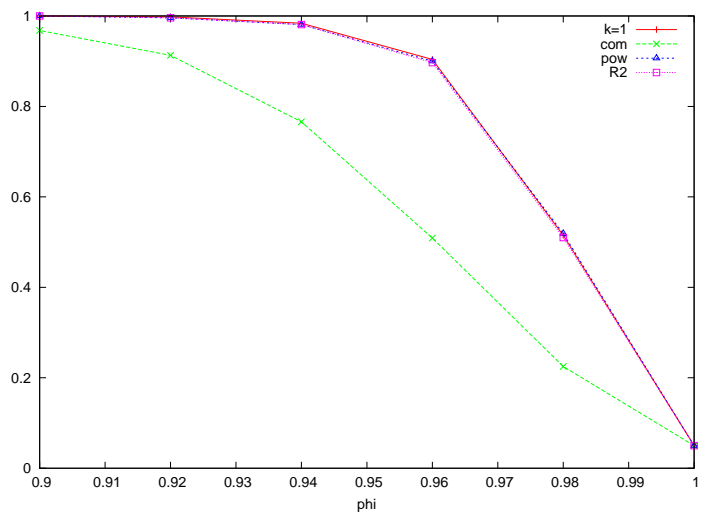
(ii-b) trend case ( $\bar{R}^2$ )

Figure 2: The performance of the proposed rules

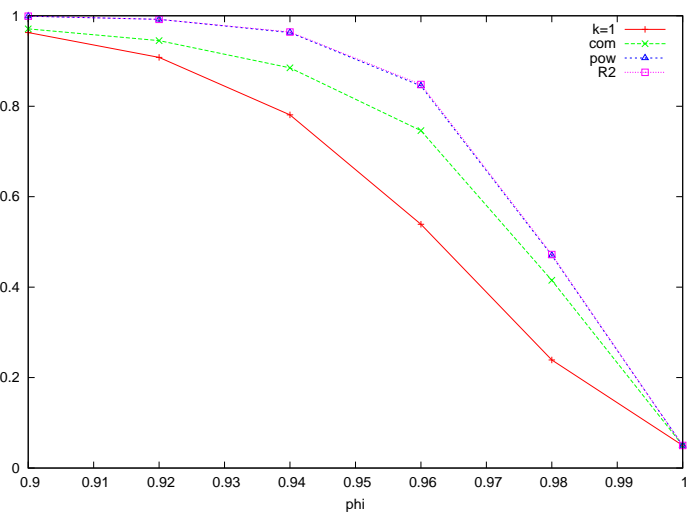




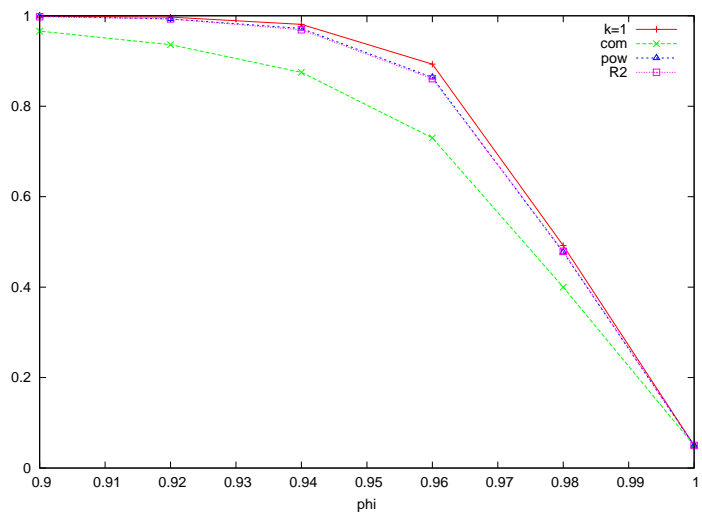
(i-a)  $N = 5$ , DGP1



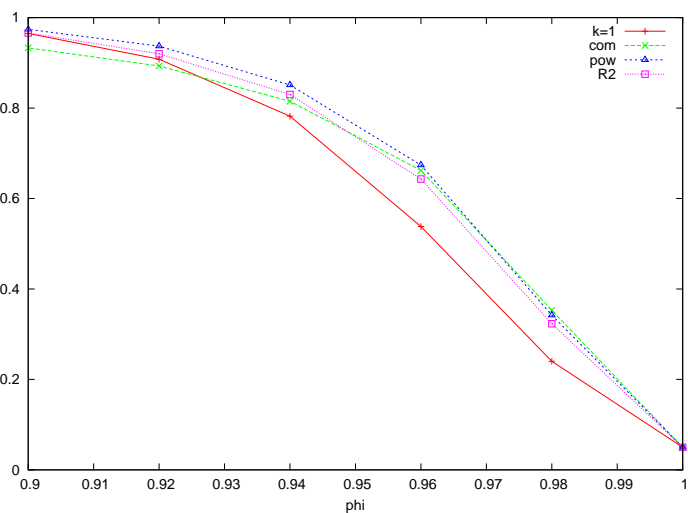
(ii-a)  $N = 5$ , DGP2



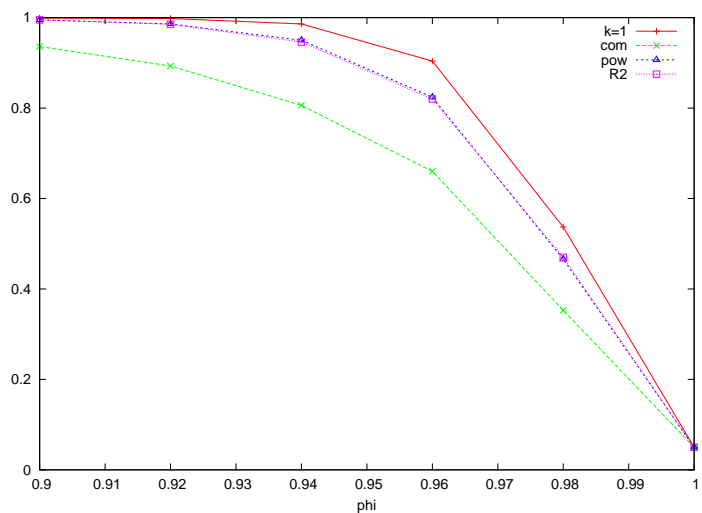
(i-b)  $N = 10$ , DGP1



(ii-b)  $N = 10$ , DGP2

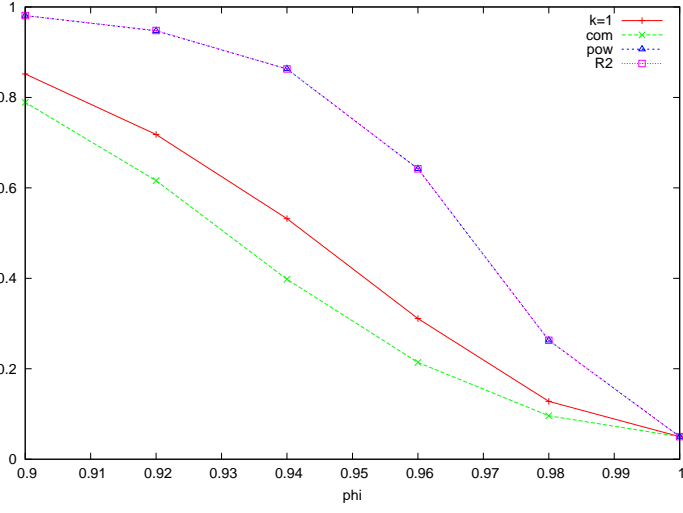


(i-c)  $N = 20$ , DGP1

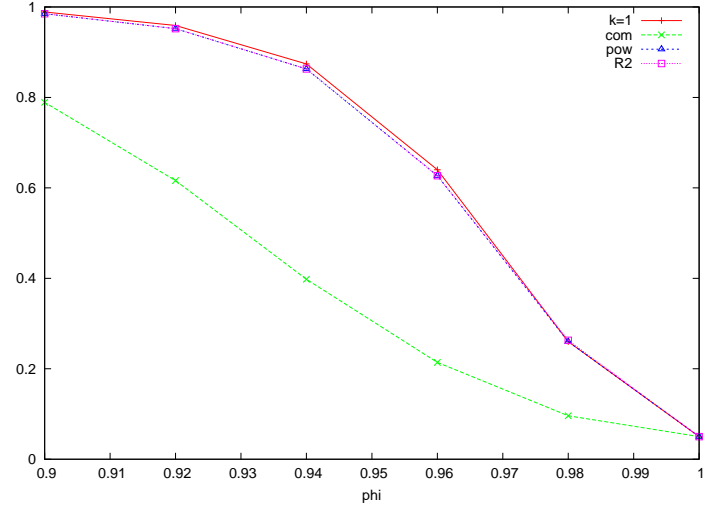


(ii-c)  $N = 20$ , DGP2

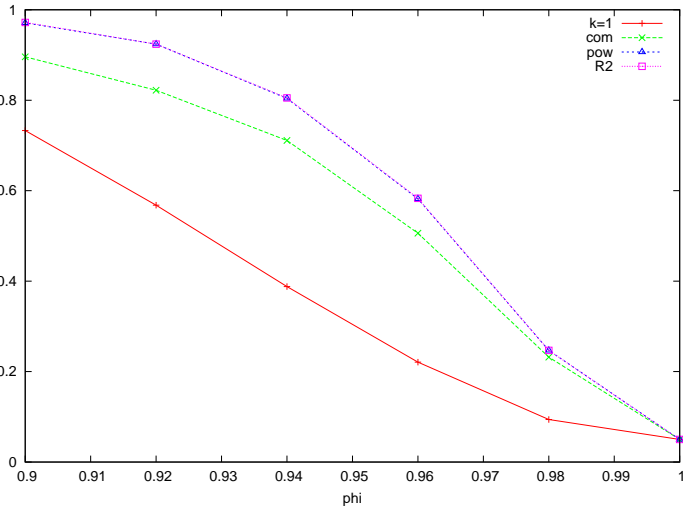
Figure 3: The size-adjusted power (constant case,  $T = 200$ ,  $R^2 = 0.6$ )



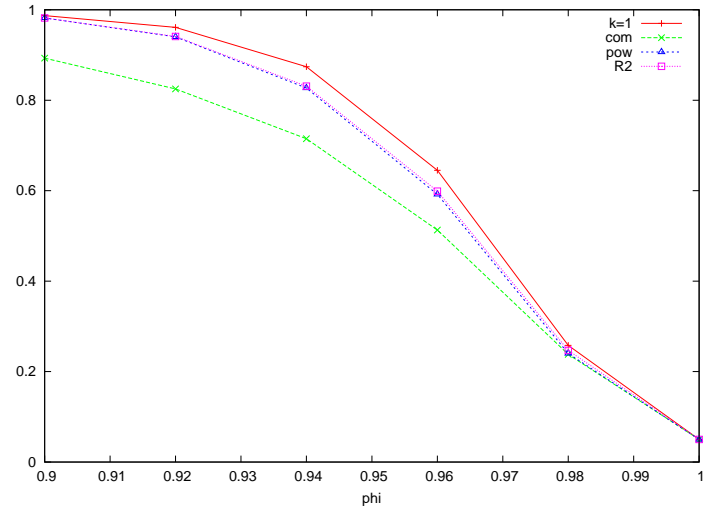
(i-a)  $N = 5$ , DGP1



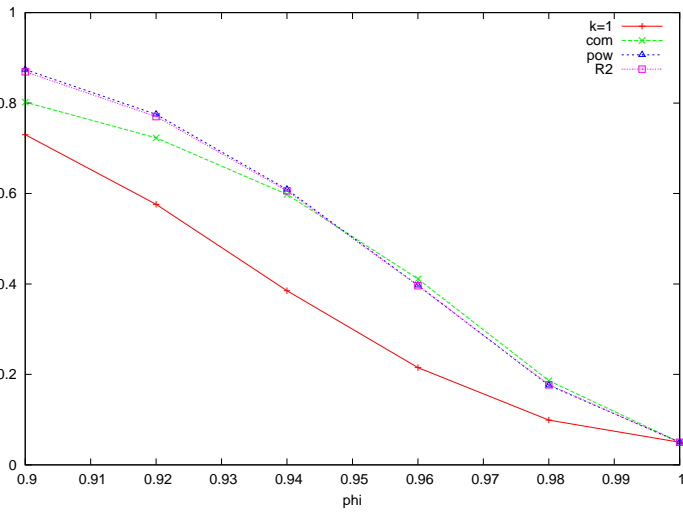
(ii-a)  $N = 5$ , DGP2



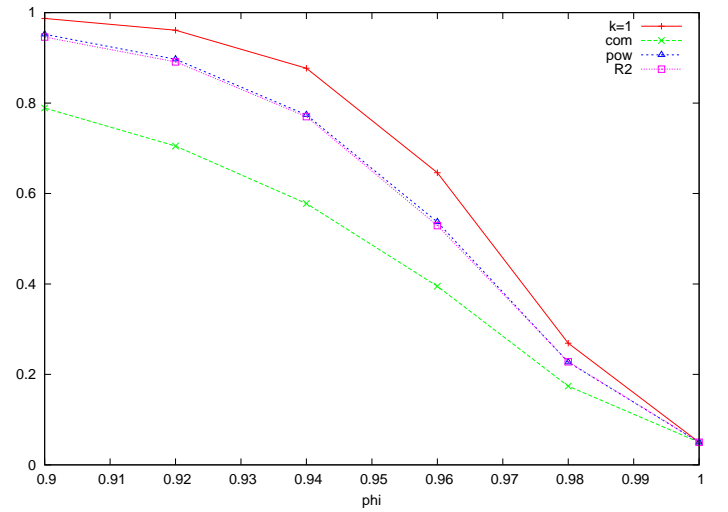
(i-b)  $N = 10$ , DGP1



(ii-b)  $N = 10$ , DGP2



(i-c)  $N = 20$ , DGP1



(ii-c)  $N = 20$ , DGP2

Figure 4: The size-adjusted power (trend case,  $T = 200$ ,  $R^2 = 0.6$ )